Tensor polyadic decomposition for antenna array processing

Souleymen Sahnoun, Pierre Comon

Tensor polyadic decomposition for antenna array processing

Souleymen Sahnoun, *Gipsa-Lab, CNRS, Grenoble France*,
Pierre Comon, *Gipsa-Lab, CNRS, Grenoble France*, `firstname.lastname@gipsa-lab.fr`

Abstract. In the present framework, a tensor is understood as a multi-way array of complex numbers indexed by three (or more) indices. The decomposition of such tensors into a sum of decomposable (i.e. rank-1) terms is called "Polyadic Decomposition" (PD), and qualified as "canonical" (CPD) if it is unique up to trivial indeterminacies. The idea is to use the CPD to identify the location of radiating sources in the far-field from several sensor subarrays, deduced from each other by a translation in space. The main difficulty of this problem is that noise is present, so that the measurement tensor must be fitted by a low-rank approximate, and that the infimum of the distance between the two is not always reached.

Our contribution is three-fold. We first propose to minimize the latter distance under a constraint ensuring the existence of the minimum. Next, we compute the Cramér-Rao bounds related to the localization problem, in which nuisance parameters are involved (namely the translations between subarrays). Then we demonstrate that the CPD-based localization algorithm performs better than ESPRIT when more than 2 subarrays are used, performances being the same for 2 subarrays. Some inaccuracies found in the literature are also pointed out.

Keywords. multi-way array ; localization ; antenna array processing ; tensor decomposition ; low-rank approximation ; complex Cramer-Rao bounds

1 Introduction

The goal is to estimate the Directions of Arrival (DoA) of R narrow-band radiating sources, which impinge on an array of sensors, formed of L identical subarrays of K sensors each. Subarrays do not need to be disjoint, but must be distinct. The hypotheses are [1, 2, 3]: (H1) sources are in the far-field, so that waves are plane; (H2) taking one subarray as reference, every subarray is deduced from the reference one by an unknown translation in space, defined by some vector $\boldsymbol{\delta}_\ell$ of $\mathbb{R}^3$, $1 < \ell \leq L$, $\boldsymbol{\delta}_1 \stackrel{\text{def}}{=} \mathbf{0}$; (H3) measurements are recorded on each sensor k of each subarray ℓ and for various time samples m, $1 \leq m \leq M$. In hypothesis (H2), the fact that translations are not exactly known is legitimate, if subarrays are arranged far away from each other, or when their location is changing with time [4]. With these hypotheses, the observation model below

can be assumed [2, 3, 4]:

$$Z(k,\ell,m) = \sum_{r=1}^{R} A_{kr}\, B_{\ell r}\, C_{mr} + N(k,\ell,m) \tag{1}$$

where $N(k,\ell,m)$ is a measurement/background noise that will be assumed normally distributed. In addition, (H4) parameters $\{A_{kr},\, 1 \le k \le K,\, 1 \le r \le R\}$ are complex numbers of unit modulus, and $B_{1r} = 1 = A_{1r},\, \forall r$. In the present framework, we shall consider subarrays that are formed of equispaced sensors, so that the following can also be assumed (H5):

$$\exists\, \psi_r : \quad A_{kr} = \exp(\jmath\pi(k-1)\cos\psi_r) \tag{2}$$

where ψ_r is the so-called Direction of Arrival (DoA) of the rth source as illustrated in Figure 1, and $\jmath = \sqrt{-1}$. Space is lacking to explain the physical context, but further details can be found in [1, 2, 3, 4]. The literature is abundant about DoA estimation, but most approaches have been based on second-order moments; see *e.g.* [5] and references therein. On the other hand, the use of space diversity via more than two subarrays is much more recent, and is due to [2]. The key originality therein is that the approach is not based on moments but proceeds by direct parameter estimation from the data.

A key ingredient in this problem is the use of complex random variables, which turn out to be very useful because the formalism is much simpler when working in baseband with complex envelopes of transmitted signals. Among the useful ingredients, we have at disposal matrix differentiation [6], Kronecker and tensor calculus [7, 8], complex differentiation and the derivation of complex Cramér-Rao bounds (see Section 3).

Notation. $\mathbb{R}$ and $\mathbb{C}$ designate the real and complex fields, respectively. Bold lower case letters, *e.g.* $\mathbf{z}$, always denote column vectors, whereas arrays with 2 indices or more are denoted by bold uppercase symbols, *e.g.* $\mathbf{V}$ or $\mathbf{Z}$. Array entries are scalar numbers and are denoted in plain font, *e.g.* z_i, V_{ij} or $Z_{k\ell m}$. The gradient of a p-dimensional function $\mathbf{f}(\mathbf{x})$ with respect to a n-dimensional variable $\mathbf{x}$ is the $p \times n$ matrix $[\partial\mathbf{f}/\partial\mathbf{x}]_{ij} = \partial f_i/\partial x_j$.

2 Tensor formalism and constrained optimization

In this framework, what is meant by *tensor* is just a multi-way array of coordinates; this is not restrictive as long as the coordinate system is fixed [8]. The noiseless part of (1) is a sum of *decomposable tensors*, whose coordinates are of the form $D_{k\ell m} = a_k\, b_\ell\, c_m$. Any tensor can be decomposed into a sum of decomposable tensors, and the minimal number of terms necessary to obtain an exact decomposition is called *tensor rank*. Hence nonzero decomposable tensors have a rank equal to 1. Because of the presence of noise, the best rank-R tensor approximate of $\mathbf{Z}$ needs to be found, for instance in the sense of the Frobenius norm, which is consistent with the log-likelihood (11) in the presence of additive Gaussian noise. However, as pointed out in [3, 8, 4] and references therein, the infimum of

$$\Upsilon(\mathbf{A},\mathbf{B},\mathbf{C}) \stackrel{\text{def}}{=} \|\mathbf{Z} - \sum_{r=1}^{R} \mathbf{a}_r \otimes \mathbf{b}_r \otimes \mathbf{c}_r\|^2$$

may not be reached. Here, $\mathbf{a}_r$, $\mathbf{b}_r$, $\mathbf{c}_r$ denote the columns of matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ defined in (1), respectively, and $\otimes$ is the tensor outer product.

It has been proved in [3] that a sufficient condition ensuring existence of the best rank-R approximation is that

$$\begin{array}{rcl} \mu_A\mu_B\mu_C & < & \dfrac{1}{R-1} \\ \text{where} \quad \mu_A = \sup_{k\neq\ell} |\mathbf{a}_k^{\mathsf{H}}\mathbf{a}_\ell|, \ \mu_B & = & \sup_{k\neq\ell} |\mathbf{b}_k^{\mathsf{H}}\mathbf{b}_\ell|, \ \mu_C = \sup_{k\neq\ell} |\mathbf{c}_k^{\mathsf{H}}\mathbf{c}_\ell|. \end{array} \tag{3}$$

Therefore, the following *differentiable* constraint, proposed in [4], can be imposed:

$$\mathcal{C}_\rho \stackrel{\text{def}}{=} 1 - R + \mu(\mathbf{A},\rho)^{-1}\mu(\mathbf{B},\rho)^{-1}\mu(\mathbf{C},\rho)^{-1} > 0, \quad \mu(\mathbf{A},\rho) \stackrel{\text{def}}{=} \left(\sum_{p<q} |\mathbf{a}_p^{\mathsf{H}}\mathbf{a}_q|^{2\rho}\right)^{1/2\rho} \tag{4}$$

In fact, the inequality between L^p norms

$$||\mathbf{x}||_\infty = \max_k\{x_k\} \leq ||\mathbf{x}||_p \stackrel{\text{def}}{=} (\sum_k x_k^p)^{1/p}, \ \forall x_k \in \mathbb{R}^+, \ p \geq 1,$$

guarantees that constraint (4) implies condition (3). In practice, the following penalized objective function has been minimized in subsequent computer simulations:

$$\Upsilon(\mathbf{A},\mathbf{B},\mathbf{C}) + \eta \, \exp(-\gamma \, \mathcal{C}_\rho(\mathbf{x})) \tag{5}$$

with $\rho = 13$, $10^{-6} \leq \eta \leq 1$, $\gamma = 5$.

3 Complex Cramér-Rao bounds

When parameters are complex, expressions of Cramér-Rao bounds (CRB) depend on the definition of the complex derivative. Since a real function is never holomorphic (unless it is constant) [9], this definition is necessary; this has been overlooked in [10]. Originally, the derivative of a real function $\mathbf{h}(\boldsymbol{\theta}) \in \mathbb{R}^p$ with respect to a complex variable $\boldsymbol{\theta} \in \mathbb{C}^n$, $\boldsymbol{\theta} = \boldsymbol{\alpha} + \jmath\boldsymbol{\beta}$, $\boldsymbol{\alpha},\boldsymbol{\beta} \in \mathbb{R}^n$, has been defined as the $p \times n$ matrix [9]:

$$\frac{\partial \mathbf{h}}{\partial \boldsymbol{\theta}} \stackrel{\text{def}}{=} \frac{\partial \mathbf{h}}{\partial \boldsymbol{\alpha}} + \jmath \frac{\partial \mathbf{h}}{\partial \boldsymbol{\beta}}$$

Even if the numerical results are independent of the definition assumed for theoretical calculations, we shall subsequently assume the definition proposed in [11], for consistency with [12]:

$$\frac{\partial \mathbf{h}}{\partial \boldsymbol{\theta}} \stackrel{\text{def}}{=} \frac{1}{2}\frac{\partial \mathbf{h}}{\partial \boldsymbol{\alpha}} - \frac{\jmath}{2}\frac{\partial \mathbf{h}}{\partial \boldsymbol{\beta}} \tag{6}$$

With this definition, one has for instance that $\partial\boldsymbol{\alpha}/\partial\boldsymbol{\theta} = \frac{1}{2}\mathbf{I}$, and $\partial\boldsymbol{\beta}/\partial\boldsymbol{\theta} = -\frac{\jmath}{2}\mathbf{I}$. This is a key difference with [9], where we had instead: $\partial\boldsymbol{\alpha}/\partial\boldsymbol{\theta} = \mathbf{I}$, and $\partial\boldsymbol{\beta}/\partial\boldsymbol{\theta} = \jmath\mathbf{I}$. Assume that parameter $\boldsymbol{\theta}$ is wished to be estimated from an observation $\mathbf{z}$, of probability distribution $\mathcal{L}(\mathbf{z};\boldsymbol{\theta})$, and denote $\mathbf{s}(\mathbf{z};\boldsymbol{\theta})$ the score function. Then we have for any function $\mathbf{h}(\boldsymbol{\theta}) \in \mathbb{R}^p$:

$$\mathsf{E}\{\mathbf{h}(\mathbf{z})\mathbf{s}(\mathbf{z};\boldsymbol{\theta})^{\mathsf{T}}\} = \frac{\partial}{\partial\boldsymbol{\theta}}\mathsf{E}\{\mathbf{h}(z)\}, \quad \text{with } \mathbf{s}(\mathbf{z};\boldsymbol{\theta})^{\mathsf{T}} \stackrel{\text{def}}{=} \frac{\partial}{\partial\boldsymbol{\theta}} \log \mathcal{L}(\mathbf{z};\boldsymbol{\theta}) \tag{7}$$

This is a direct consequence of the fact that $\mathsf{E}\{\mathbf{s}\} = \mathbf{0}$, valid if derivation with respect to $\boldsymbol{\theta}$ and integration with respect to $\Re(\mathbf{z})$ and $\Im(\mathbf{z})$ can be permuted. Now let $\mathbf{t}(\mathbf{z})$ be an unbiased estimator of $\boldsymbol{\theta}$. Then, following [9], one can prove that $\mathsf{E}\{\mathbf{t}\,\mathbf{s}^{\mathsf{T}}\} = \mathsf{E}\{(\mathbf{t}-\boldsymbol{\theta})\,\mathbf{s}^{\mathsf{T}}\} = \mathbf{I}$ and $\mathsf{E}\{\mathbf{t}\,\mathbf{s}^{\mathsf{H}}\} = \mathbf{0}$. Finally, by expanding the covariance matrix of the random vector $(\mathbf{t}-\boldsymbol{\theta}) - \mathbf{F}^{-1}\mathbf{s}^*$, one readily obtains that:

$$\mathbf{V} \geq \mathbf{F}^{-1}, \quad \text{with } \mathbf{V} \stackrel{\text{def}}{=} \mathsf{E}\{(\mathbf{t}-\boldsymbol{\theta})(\mathbf{t}-\boldsymbol{\theta})^{\mathsf{H}}\} \text{ and } \mathbf{F} \stackrel{\text{def}}{=} \mathsf{E}\{\mathbf{s}^*\,\mathbf{s}^{\mathsf{T}}\} \tag{8}$$

Note that the definition of the Fisher information matrix is the complex conjugate of that of [9], because of a different definition of the complex derivation (and hence a different definition of the complex score function).

4 Cramér-Rao bounds of the localization problem

In the presence of R sources, the observations can be stored in a three-way array unfolded in vector form:

$$\mathbf{z} = \sum_{r=1}^{R} \mathbf{a}_r \boxtimes \mathbf{b}_r \boxtimes \mathbf{c}_r + \mathbf{n}, \quad \mathbf{z} \in \mathbb{C}^{KLM} \tag{9}$$

where $\boxtimes$ denotes the Kronecker product, and the additive noise $\mathbf{n}$ is assumed to follow a circularly-symmetric complex normal distribution. Let

$$\boldsymbol{\theta} = [\underbrace{\psi_1, \ldots, \psi_R}_{\boldsymbol{\psi}}, \underbrace{\bar{\mathbf{b}}_1^{\mathsf{T}}, \ldots, \bar{\mathbf{b}}_R^{\mathsf{T}}, \mathbf{c}_1^{\mathsf{T}}, \ldots, \mathbf{c}_R^{\mathsf{T}}}_{\boldsymbol{\xi}}, \underbrace{\bar{\mathbf{b}}_1^{\mathsf{H}}, \ldots, \mathbf{c}_R^{\mathsf{H}}}_{\boldsymbol{\xi}^*}] \tag{10}$$

denote the unknown parameter vector, where $\bar{\mathbf{b}}_r \stackrel{\text{def}}{=} [B_{2,r}, \ldots, B_{L,r}]^{\mathsf{T}}$. It is useful to include both $\boldsymbol{\xi}$ and $\boldsymbol{\xi}^*$, in case the distribution of the estimate $\hat{\boldsymbol{\xi}}$ is not circularly-symmetric, *i.e.* $\mathsf{E}\{\hat{\boldsymbol{\xi}}\,\hat{\boldsymbol{\xi}}^{\mathsf{T}}\} \neq \mathbf{0}$. The aim here is to derive the CRB of the parameters in $\boldsymbol{\theta}$. The CRB for factor matrices have been computed in [12]. However, it should be emphasized that, unlike [12], no assumption is needed on the elements of matrix $\mathbf{C}$ to derive the CRB. In fact, assuming that the first row of $\mathbf{A}$ and $\mathbf{B}$ is fixed to $[1, \ldots, 1]_{1\times R}$ is sufficient. The latter assumption is satisfied in the considered array configuration (hypothesis: H4). The log-likelihood then takes the form:

$$\log \mathcal{L}(\mathbf{z}, \boldsymbol{\theta}) = -KLM \log(\sigma^2 \pi) - \frac{1}{\sigma^2}(\mathbf{z} - \boldsymbol{\mu}(\boldsymbol{\theta}))^{\mathsf{H}}(\mathbf{z} - \boldsymbol{\mu}(\boldsymbol{\theta})) \tag{11}$$

where $\boldsymbol{\mu}(\boldsymbol{\theta})$ is the noise free part of $\mathbf{z}$. The CRB for unbiased estimation of the complex parameters $\boldsymbol{\theta}$ is equal to the inverse of the Fisher information matrix $\mathbf{F}$, defined in equation (8). Then, a straightforward calculation yields:

$$\mathbf{s}^{\mathsf{T}} = \frac{1}{\sigma^2}\left[\mathbf{n}^{\mathsf{T}}\frac{\partial \boldsymbol{\mu}^*}{\partial \boldsymbol{\theta}} + \mathbf{n}^{\mathsf{H}}\frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\theta}}\right] \tag{12}$$

where $\mathbf{n} = \mathbf{z} - \boldsymbol{\mu}$. By substituting the score function $\mathbf{s}$ by its expression, and since $\mathsf{E}\left\{\mathbf{n}\mathbf{n}^{\mathsf{H}}\right\} = \sigma^2\mathbf{I}_{KLM}$ and $\mathsf{E}\left\{\mathbf{n}\mathbf{n}^{\mathsf{T}}\right\} = \mathbf{0}$, the Fisher information matrix can be written as:

$$\mathbf{F} = \frac{1}{\sigma^2}\left[\left(\frac{\partial \boldsymbol{\mu}^*}{\partial \boldsymbol{\theta}}\right)^{\mathsf{H}}\left(\frac{\partial \boldsymbol{\mu}^*}{\partial \boldsymbol{\theta}}\right) + \left(\frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\theta}}\right)^{\mathsf{H}}\left(\frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\theta}}\right)\right] \tag{13}$$

Since parameters in $\boldsymbol{\psi}$ are real and those in $\boldsymbol{\xi}$ are complex, a first writing of the derivatives in (13) is:

$$\frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\theta}} = \left[\frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\psi}}, \quad \frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\xi}}, \quad \mathbf{0}\right] \quad \text{and} \quad \frac{\partial \boldsymbol{\mu}^*}{\partial \boldsymbol{\theta}} = \left[\left(\frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\psi}}\right)^*, \quad \mathbf{0}\,, \quad \left(\frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\xi}}\right)^*\right] \tag{14}$$

Therefore, the Fisher information matrix becomes:

$$\mathbf{F} = \frac{1}{\sigma^2}\begin{bmatrix} 2\,\mathrm{Re}\,\{\mathbf{G}_{11}\} & \mathbf{G}_{12} & \mathbf{G}_{12}^* \\ \mathbf{G}_{12}^{\mathsf{H}} & \mathbf{G}_{22} & \mathbf{0} \\ \mathbf{G}_{12}^{\mathsf{T}} & \mathbf{0} & \mathbf{G}_{22}^* \end{bmatrix} \tag{15}$$

$$\text{where} \quad \mathbf{G}_{ij} = \left(\frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\theta}_i}\right)^{\mathsf{H}} \left(\frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\theta}_j}\right), \ (i,j) \in \{1,2\} \times \{1,2\}, \ \boldsymbol{\theta}_1 = \boldsymbol{\psi} \ \text{ and } \boldsymbol{\theta}_2 = \boldsymbol{\xi}. \tag{16}$$

In view of (15), it is clear that the introduction of $\boldsymbol{\xi}^*$ in the parameter vector was not necessary. With a non circular complex gaussian noise, this would not have been the case. To complete the calculation of $\mathbf{F}$, it remains to give partial derivative expressions of $\boldsymbol{\mu}$ with respect to $\boldsymbol{\psi}$ and $\boldsymbol{\xi}$.

Derivatives of $\boldsymbol{\mu}$ with respect to $\boldsymbol{\psi}$

Using the chain rule we have

$$\frac{\partial \boldsymbol{\mu}}{\partial \psi_f} = \left(\frac{\partial \boldsymbol{\mu}}{\partial\, \mathbf{a}_f^{\mathsf{T}}}\right)\left(\frac{\partial \mathbf{a}_f^{\mathsf{T}}}{\partial \psi_f}\right) \tag{17}$$

and $[\partial \boldsymbol{\mu}/\partial \mathbf{a}_f^{\mathsf{T}}]$ can be computed using complex derivative formulas. Then, we obtain:

$$\frac{\partial \boldsymbol{\mu}}{\partial \mathbf{a}_f^{\mathsf{T}}} = \mathbf{I}_K \boxtimes \mathbf{b}_f \boxtimes \mathbf{c}_f \quad \in \mathbb{C}^{KLM \times K}, \ 1 \le f \le R. \tag{18}$$

To calculate $[\partial \mathbf{a}_f^{\mathsf{T}}/\partial \psi_f]$, we use the expressions of the considered sensor array configuration, namely equation (2), which yields:

$$\frac{\partial \mathbf{a}_f^{\mathsf{T}}}{\partial \psi_f} = -\jmath\pi \sin\psi_f \,(\mathbf{a}_f \boxdot \mathbf{v}_K) \tag{19}$$

where $\mathbf{v}_K = [0, 1, \ldots, K-1]^{\mathsf{T}}$. By substituting (18) and (19) in (17), we get

$$\frac{\partial \boldsymbol{\mu}}{\partial \psi_f} = -\jmath\pi \sin\psi_f \ (\mathbf{I}_K \boxtimes \mathbf{b}_f \boxtimes \mathbf{c}_f)\,(\mathbf{a}_f \boxdot \mathbf{v}_K) \stackrel{\text{def}}{=} \boldsymbol{\phi}_{\psi_f} \tag{20}$$

$$\text{and} \quad \frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\psi}} = [\boldsymbol{\phi}_{\psi_1}, \ldots, \boldsymbol{\phi}_{\psi_R}] \in \mathbb{C}^{KLM \times R} \tag{21}$$

Derivatives of $\boldsymbol{\mu}$ with respect to $\boldsymbol{\xi}$

Taking partial derivatives of $\boldsymbol{\mu}$ with respect to $\bar{\mathbf{b}}_f^\mathsf{T}$ and $\mathbf{c}_f^\mathsf{T}$, we obtain:

$$\frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \bar{\mathbf{b}}_f^\mathsf{T}} = (\mathbf{a}_f \boxtimes \mathbf{I}_{LM})(\mathbf{I}_L \boxtimes \mathbf{c}_f)\mathbf{J}_L \overset{\text{def}}{=} \boldsymbol{\phi}_{\bar{\mathbf{b}}_f} \quad \in \mathbb{C}^{KLM\times(L-1)} \tag{22}$$

$$\frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \mathbf{c}_f^\mathsf{T}} = \mathbf{a}_f \boxtimes \mathbf{b}_f \boxtimes \mathbf{I}_M \overset{\text{def}}{=} \boldsymbol{\phi}_{\mathbf{c}_f} \quad \in \mathbb{C}^{KLM\times M} \tag{23}$$

where $\mathbf{J}_L = [\mathbf{0}_{(L-1),1}\ \mathbf{I}_{L-1}]^\mathsf{T} \in \mathbb{C}^{L\times(L-1)}$ is a selection matrix. To sum up,

$$\frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\xi}} = [\boldsymbol{\phi}_{\bar{\mathbf{b}}_1}, \ldots, \boldsymbol{\phi}_{\bar{\mathbf{b}}_R}, \boldsymbol{\phi}_{\mathbf{c}_1}, \ldots, \boldsymbol{\phi}_{\mathbf{c}_R}] \in \mathbb{C}^{KLM\times R(L+M-1)} \tag{24}$$

DoA Cramér-Rao bound

The CRB related to DoAs only is obtained as the first leading $R \times R$ block in matrix $\mathbf{F}^{-1}$, where $\mathbf{F}$ is defined in (15). Doing this assumes that translations $\boldsymbol{\delta}_\ell$ are nuisance parameters, *i.e.* unknown but not of interest. This realistic context has been overlooked in the literature.

5 Computer results

To evaluate the efficiency of the proposed method, we compare its performances to two other algorithms, ESPRIT and MUSIC [13, 14]. The performance criterion is the *total* mean square error (total MSE) of the DoA: $\frac{1}{RN}\sum_{r=1}^{R}\sum_{n=1}^{N}(\hat{\psi}_{r,n} - \psi_r)^2$ where $\hat{\psi}_{r,n}$ is the estimated DoA at the n-th Monte-Carlo trial and N is the number of trials. The deterministic CRB computed in the previous section is reported as a benchmark. The considered scenario on which the proposed algorithm is tested can be of interest in numerous applications, where translations $\boldsymbol{\delta}_\ell$ are unknown. Note that the CRB of the DoA where locations of all sensors are known can be found in [13, 14]. The three examples we study in this section are reported in the table below:

	Subarrays	Translations	DoA
Example 1	$L = 2$	$\boldsymbol{\delta}_2 = [0, 25\lambda, 0]^\mathsf{T}$	$40°, 64°, 83°$
Example 2	$L = 3$	$\boldsymbol{\delta}_2 = [0, 25\lambda, 0]^\mathsf{T}$, $\boldsymbol{\delta}_3 = [0, 37.5\lambda, 5\lambda]^\mathsf{T}$	$40°, 64°, 83°$
Example 3	$L = 3$	$\boldsymbol{\delta}_2 = [0, 25\lambda, 0]^\mathsf{T}$, $\boldsymbol{\delta}_3 = [0, 37.5\lambda, 5\lambda]^\mathsf{T}$	$7°, 64°, 83°$

where $\lambda = \omega/2\pi\varsigma$ is the wavelength and ς the wave celerity. In all examples, each subarray is an ULA array of 4-element with half-wavelength spacing (see Figure 1), and the narrowband source signals have the same power.

In all experiments, $M = 200$ time samples are used, and 200 Monte-Carlo simulations are run for each SNR level. Figures 2, 3 and 4 report the MSE of the DoA obtained in examples 1, 2 and 3, respectively.

Example 1. This experiment shows that: (i) the proposed CP algorithm exhibits the same performances as ESPRIT, which makes sense, (ii) MUSIC performs the best, but exploits more information, namely the exact knowledge of sensor locations, whereas this information is actually not available in the present scenario. Hence MUSIC performances just serve as a reference.

Example 2. This experiment shows that the proposed algorithm yields better results than EPSRIT. The reason is that ESPRIT uses at most two subarrays, whereas the proposed algorithm uses all of them. Again, MUSIC is reported just as a reference benchmark.

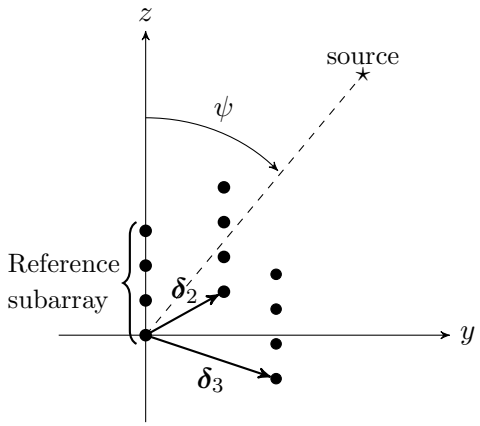

Figure 1: One source ($R = 1$) radiating on a sensor array with $L = 3$ subarrays.

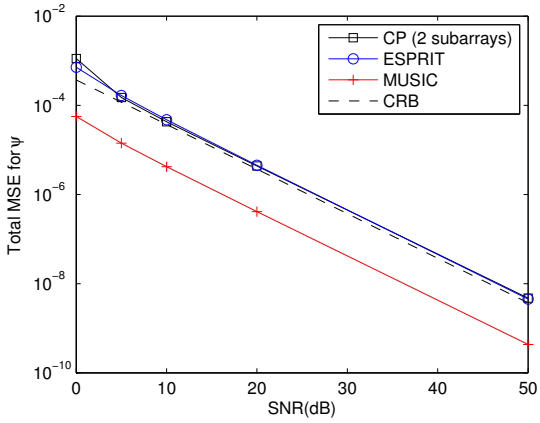

Figure 2: Total DoA error versus SNR, with $L = 2$ subarrays, $\boldsymbol{\psi} = [40°, 65°, 83°]$.

Example 3. This experiment shows the same results as in example 2, except for an increase in MSE at low SNR, which is due to the direction of arrival $\psi = 7°$. Actually, for an ULA, the source localization accuracy degrades as the DoA come closer to the end-fire, so that the so-called *threshold region* (which always exists at low SNR) becomes visible.

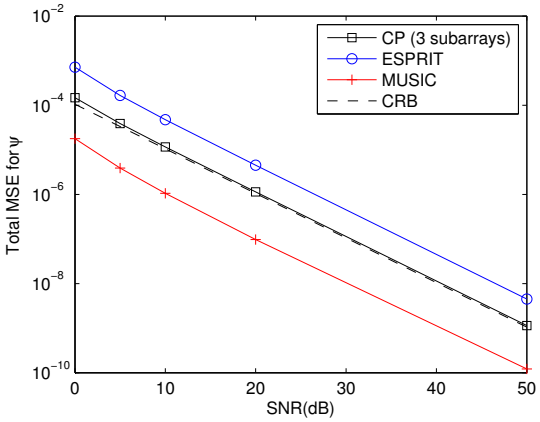

Figure 3: Total DoA error versus SNR, with $L = 3$ subarrays, $\boldsymbol{\psi} = [40°, 65°, 83°]$.

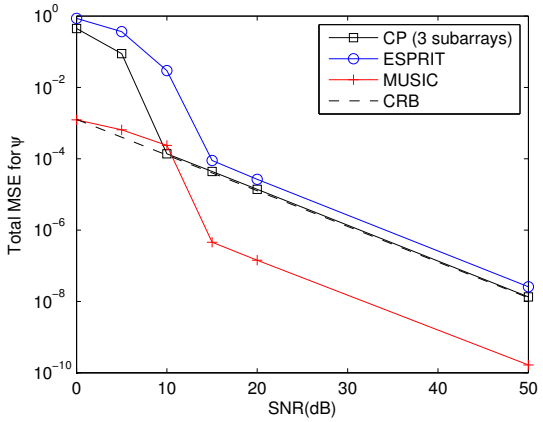

Figure 4: Total DoA error versus SNR, with $L = 3$ subarrays, $\boldsymbol{\psi} = [7°, 65°, 83°]$.

6 Conclusion

The source localization problem is taken as an illustration of the interest in resorting to CP tensor decomposition. We took the opportunity of this illustration to emphasize the usefulness of complex formalism when computing the CRB. Our contributions include the computation of CRB of DoAs when space translations are unknown (section 4), and an original algorithm to compute the CP decomposition under a constraint ensuring the existence of the best low-rank approximate (section 2).

Some inaccuracies on this subject may be found in the literature: (i) in [10], functions of the complex variable are assumed holomorphic, whereas real functions never are; (ii) in [12], CRB are derived, but without assuming that factor matrix **A** is parameterized by angles of arrival; moreover, additional constraints have been added therein to fix permutation ambiguities, whereas they are not necessary; (iii) in [14], CRB are computed for the ESPRIT technique, but translations are assumed known whereas they are actually unknown nuisance parameters; if they are known, ESPRIT cannot perform better than MUSIC.

Acknowledgement This work has been funded by the European Research Council under the 7th Framework Programme FP7/2007–2013 Grant Agreement no. 320594.

Bibliography

[1] R. Roy and T. Kailath, "ESPRIT - estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust. Speech Signal Proc.*, vol. 37, pp. 984–995, July 1989.

[2] N. D. Sidiropoulos, R. Bro, and G. B. Giannakis, "Parallel factor analysis in sensor array processing," *IEEE Trans. Sig. Proc.*, vol. 48, no. 8, pp. 2377–2388, Aug. 2000.

[3] L.-H. Lim and P. Comon, "Blind multilinear identification," *IEEE Trans. Inf. Theory*, vol. 60, no. 2, pp. 1260–1280, Feb. 2014, open access.

[4] S. Sahnoun and P. Comon, "Deterministic blind identification in antenna array processing," in *8th IEEE SAM Workshop*, A Coruna, Spain, June 22-25 2014, hal-00957357.

[5] H. Krim and M. Viberg, "Two decades of array signal processing research," *IEEE Sig. Proc. Mag.*, pp. 67–95, July 1996.

[6] W. J. Vetter, "Derivative operations on matrices," *IEEE Trans. Auto. Control*, vol. 15, no. 2, pp. 241–244, 1970.

[7] J. W. Brewer, "Kronecker products and matrix calculus in system theory," *IEEE Trans. on Circuits and Systems*, vol. 25, no. 9, pp. 114–122, Sept. 1978.

[8] P. Comon, "Tensors: a brief introduction," *IEEE Sig. Proc. Magazine*, vol. 31, no. 3, May 2014, special issue on BSS. hal-00923279.

[9] P. Comon, "Estimation multivariable complexe," *Traitement du Signal*, vol. 3, no. 2, pp. 97–101, Apr. 1986, hal-00979476.

[10] A. van den Bos, "A Cramér-Rao bound for complex parameters," *IEEE Trans. Sig. Proc.*, vol. 42, no. 10, pp. 2859, Oct. 1994.

[11] A. Hjørungnes and D. Gesbert, "Complex-valued matrix differentiation: Techniques and key results," *IEEE Trans. Sig. Proc.*, vol. 55, no. 6, pp. 2740–2746, June 2007.

[12] X. Liu and N. Sidiropoulos, "Cramér-Rao lower bounds for low-rank decomposition of multidimensional arrays," *IEEE Trans. Sig. Proc.*, vol. 49, no. 9, pp. 2074–2086, 2001.

[13] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood, and Cramer-Rao bound," *IEEE Trans. Acoust. Speech Sig. Proc.*, vol. 37, no. 5, pp. 720–741, 1989.

[14] B. Ottersten, M. Viberg, and T. Kailath, "Performance analysis of the total least squares ESPRIT algorithm," *IEEE Trans. Sig. Proc.*, vol. 39, no. 5, pp. 1122–1135, 1991.